DOCUMENT RESUME

ED 465 211 EA 031 691

AUTHOR Goertz, Margaret E.
TITLE The Federal Role in Defining "Adequate Yearly Progress": The Flexibility/Accountability Trade-Off.
INSTITUTION Consortium for Policy Research in Education, Philadelphia, PA.
SPONS AGENCY National Inst. on Educational Governance, Finance, Policymaking, and Management (ED/OERI), Washington, DC.
PUB DATE 2001-00-00
NOTE 16p.
AVAILABLE FROM For full text: http://www.cpre.org/Publications/Publications_Research.htm.
PUB TYPE Reports - Evaluative (142)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Accountability; Elementary Secondary Education; Federal Government; *Federal Legislation; Federal Regulation; *Government Role; Program Effectiveness; Program Evaluation

ABSTRACT

Bills passed by the U.S. House of Representatives and the U.S. Senate in spring 2001 to reauthorize the Elementary and Secondary Education Act (ESEA) reinforce and strengthen the education accountability provisions contained in the Improving America's Schools Act (IASA) of 1994. This paper evaluates the accountability provisions in this legislation against the goals of equity, feasibility, flexibility, and focus. The research for this work is based on state accountability systems and hypothetical examples that illustrate the operation of the legislative provisions. Research questions considered include whether the proposed system of provisions is equitable, feasible, flexible, and focused. Recommendations include: state accountability systems should set improvement goals that ensure that all students meet their state's proficient level of achievement within an appropriate timeframe; adequate yearly progress (AYP) definitions should ensure continuous progress, yet be flexible; the AYP should focus on closing the achievement gap; accountability timelines should create incentives to improve practice, yet ensure that schools have sufficient time and resources to make lasting changes; and schools and school districts should have sufficient resources to help all students reach their state's standards. (Contains 13 references.) (RT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED 465 211

The Federal Role in Defining "Adequate Yearly Progress:" The Flexibility/Accountability Trade-off

Margaret E. Goertz[1]
Consortium for Policy Research in Education
Graduate School of Education
University of Pennsylvania

Bills passed by the U. S. House of Representatives and the U.S. Senate in spring 2001 to reauthorize the Elementary and Secondary Education Act (ESEA) not only reinforce but strengthen the education accountability provisions contained in the Improving America's Schools Act (IASA) of 1994. Title I of the IASA requires states to establish challenging content and performance standards, implement assessments that measure student performance against these standards, and hold schools and school systems accountable for the achievement of *all* students. Specifically, states and local school districts must determine whether schools and school districts are making adequate yearly progress (AYP) in bringing students up to state standards, identify for school improvement any school that does not made adequate yearly progress for two consecutive school years, target resources to these schools, and, if necessary, take corrective action. The IASA calls for AYP to be defined

> in a manner that (1) results in continuous and substantial yearly improvement of each school and local education agency sufficient to achieve the goal of all children…meeting the state's proficient and advanced levels of achievement; [and] (2) is sufficiently rigorous to achieve that goal within an appropriate timeframe (As cited in Elmore and Rothman, 1999, p. 85).

Provisions in HR1 (the "No Child Left Behind Act of 2001") and S1 (the "Better Education for Students and Teachers Act") make several critical changes to the IASA. Generally, these bills:

- Expand required testing in reading and mathematics from at least once during each of three grade spans—third-to-fifth, sixth-to-ninth, and tenth-to-twelfth—to every student, annually, in grades three through eight.
- Require states to measure the progress of *all* schools and school districts using the same definition of adequate yearly progress.

---

[1] The opinions expressed in this paper are those of the author and do not reflect the views of the Consortium for Policy Research in Education, its institutional members, or the U.S. Department of Education. The author thanks the Center on Education Policy for allowing CPRE to publish this paper on its web site.

EA031691

BEST COPY AVAILABLE

- Establish a target date by which all students must score at a state-defined "proficient" level.
- Specify how states should measure adequate yearly progress.
- Delineate the types of corrective action that states and districts should take when schools fail to improve.
- Hold schools, school districts and states accountable for the yearly progress of subgroups of students; and
- Shorten the timeline for identifying schools for improvement, allowing public school choice, taking corrective action, and restructuring schools.

President Bush proposed and Congress enacted many of these changes in response to what researchers, policymakers and advocacy groups identified as shortcomings in the implementation of the IASA's accountability provisions. While the intent of the IASA was to create single and "seamless" accountability systems that would treat all schools equally, only 22 states had single accountability systems in place by 2000-2001. More than half of the states had dual accountability systems where Title I schools were subject to different measures of adequate yearly progress. States also differ in the percentage of students that schools are expected to bring up to their "proficient" standard, timelines for meeting these performance goals (if they have timelines), and the extent to which their Title I AYP definitions include measures of continuous progress. Only one-third of the states focus on closing the gap between low- and high-achieving students in their Title I schools, and few states require schools to close achievement gaps between White students and students of color, between economically disadvantaged and economically advantaged students, between students with and without disabilities, or between English-dominant and limited English-speaking students (Goertz and Duffy, 2001; NATI/IRP, 2001; U.S. Department of Education, 2001).

Yet some members of the policy, research and advocacy communities have raised questions about the potential impact of the more prescriptive accountability provisions of HR1 and S1. Of particular concern are the bills' definitions of AYP and the timelines for getting all students and subgroups of students to the "proficient" standard and for identifying and sanctioning low-performing schools. In this paper, I evaluate these provisions against the goals of equity, feasibility, flexibility, and focus. I chose not to simulate the actual impact of the provisions on schools in any particular state.[2] Rather, I base my analysis on research I have conducted on state accountability systems (e.g., Goertz and Duffy, 2001) and hypothetical

[2] I did not conduct these simulations for two reasons. First, very few states report on their web sites the kinds of data needed to do these analyses—the number and percent of students who score at the proficient level, by race/ethnicity and poverty, in each school. Those states that do report these data (such as Massachusetts and Texas) have established different performance standards that make comparing the impact of the federal provisions difficult. Second, I was not sure how to operationalize the Senate's AYP definition, particularly how to weight the lowest-performing groups and/or the groups making the greatest improvement.

examples that illustrate the operation of the legislative provisions. I end with a set of recommendations for federal policymakers that are designed to meet the intent of HR1 and S1, while addressing some of the limitations of the proposed provisions.

Six assumptions underlie my analysis.

1. The goal of Title I is to ensure high standards for all students and to help all students meet their state's proficient (and advanced) levels of achievement within an appropriate time frame (IASA, 1994).

2. States, local school districts and schools should be given maximum flexibility in the design of policy in exchange for greater accountability for the continuous progress of all students and the closing of achievement gaps.

3. To ensure that all students meet high standards, states and localities must include disaggregated student achievement data in their accountability systems.

4. Measures of adequate yearly progress must be statistically valid and reliable, and minimize the probability that a school will be misclassified.

5. A well-designed accountability system uses a mix of incentives—ranging from public reporting to school restructuring—to drive school change. The types of incentives applied to any school should be sensitive to that school's demographics, level and range of student performance, and school capacity.

6. States and school districts have limited human and fiscal resources to support school improvement efforts.

Is the Proposed System Equitable?

The House and Senate bills take three steps to enhance the equitable treatment of schools and students--single accountability systems, expanded reporting on subgroup performance, and the inclusion of subgroup performance in state accountability systems.

Single accountability systems

It is important for states to use the same measures and set the same expectations for all schools under their jurisdiction. Both the House and Senate bills require states to enact a single accountability system for all schools and school districts, at least up to the point of identifying schools for improvement. States that currently operate dual accountability systems may object, however, to having the federal government dictate the substance of their accountability policies if it affects non-Title I schools.

Disaggregation of data for reporting

The House and Senate bills tighten and expand the reporting requirements of IASA. Under current law, the results of final state assessments must be disaggregated for the state, districts and schools by six student groups: gender, major racial/ethnic group, English-

proficiency status, disabled versus non-disabled status, migrant status and economically disadvantaged versus non-economically disadvantaged status. States must also provide for reporting these disaggregated data if they are statistically sound, but public reporting of these data is uneven across the states. Both the House and Senate bills require that starting in 2003-2004, states and local districts issue and disseminate state, local and school report cards that include disaggregated information on student performance on state assessments, graduation and dropout rates; whether a school is identified for school improvement; a comparison of students' scores with students in the state as a whole; and the professional qualifications of teachers. Report cards under the House bill must also compare the actual achievement of each group of students to their AYP objectives.

Disaggregation of data for accountability

The House and Senate bills also require states to establish separate annual objectives for student subgroups (other than by gender and migrant status). The bills differ significantly, however, in how these AYP objectives are set and how they can be met.

The House bill requires states to set the same progress measures for subgroups as for the total student population. This approach has the advantage of focusing attention on the continuous progress of all groups of students. In practice, however, the targets for some groups of students (e.g., students of color, economically disadvantaged students, LEP students, etc.) will generally be higher than those for their comparison groups, because they generally start at a lower level of performance. For example, State A has chosen to use a "relative growth" model of AYP, the second option available under HR1. The minimum rate of increase for all students and for subgroups of students is the difference between 100% and the actual percent of students at the proficient level, divided by the number of years allowed to reach the target. As shown in Table 1, the annual AYP measure for School 1 overall is 5.6% when given 10 years to reach its target. The annual growth rate is 4.4% for White students, 7.3% for African American students, and 6.3% for Hispanic students. These annual rates are lower across the board when the target date is set at 12 years out. Thus, the annual goals for the school and for each major subgroup are a function of two factors: (1) the percentage of students scoring below proficient in a given year, and (2) the number of years schools have to reach the target of remaining before schools must get to 100% proficient. [3] The calculation of AYP is not affected by the demographic composition

[3] This is not the case under the "absolute target" approach, the other option in HR1. Here, the state must set the same absolute goal for each subgroup (e.g., 35% of all students and all subgroups meeting proficiency in Year 1) and raise this goal each year until it reaches 100%. Under this method, schools may have to focus more on raising the performance of minority students in the early years until they are on performing at levels close to the White students in their school. At that time, each group will be expected to show equal progress. Assume, for example, that a state sets its Year 1 proficiency target at 35%, and plans to raise this target by 7 percentage points a year ([100%-35%]/9 years). In Year 1, only the African American students in School 1 fall short of this goal. In year 2, Hispanic students will fall short of the new goal of 42% (35%+7%). It isn't until Year 5 that White students fail to meet the target (which by then is 63%).

of a school, however. That is, the annual goals for each subgroup will be the same in a racially- and/or economically-diverse school as in a more homogeneous building.[4]

The AYP formula in the Senate bill gives additional weight to those subgroups that are furthest from the goal of 100% proficiency, but does not hold schools accountable for meeting subgroup objectives beyond a minimum growth rate of one percent. If schools have a significant proportion of White or economically advantaged students, they can meet the Senate's AYP goals without increasing the performance of minority or disadvantaged students significantly (Education Trust, 2001). For example, assume that the AYP for School 1 rises from 5.5% to 6.1% when the weights in the Senate bill are applied. The school can meet its target if 61 additional students attain proficiency, regardless of their race. The goal can be met by raising the percent of White students who are proficient by 8% (40 students), the percent of African American students by 4% (10 students) and the percent of Hispanic students by 4% (10 students). At the end of the year, the achievement gap will have widened by 4% between the White and minority students. In addition, at this rate of change, it will take 16 to 18 years before the minority students reach 100% proficiency. If one applies the minimum growth rate of one percent, it could take 60 to 70 years to bring everyone up to a proficient level.

Is the Proposed System Feasible?

The accountability provisions in both bills raise a number of technical, political and educational feasibility issues.

Technical feasibility

The House bill requires states to set annual improvement goals for all students and subgroups of students that will result in all students scoring at their state's proficient level in 12 years. Both AYP formulas—setting an escalating "absolute target" or a "relative growth" target—require schools to show fairly linear growth in student performance annually. Recent analyses of accountability systems in North Carolina (Kane and Staiger, 2001) and California (Hill, 2000) raise serious questions, however, about the statistical reliability of annual, observed changes in school performance in general, and for subgroups in particular. As Hill (2000) writes

> The performance of any school on any index will vary from year to year, in part because of systematic changes, but also because of random fluctuations. A school that scored higher than another in one year might score lower in another year, not because of any fundamental change in teaching practices or student population, but because the sample of students, and their performance on the given day that testing was done, varied (p. 1).

Hill found that even when the true growth in California schools is nearly equal to their growth target, they may have only a 50% chance of having their *observed* scores increase by as much as their growth target. Kane and Staiger report that sampling variation and other non-

[4] The student body composition does affect the AYP goal for the entire school. A school with a larger number of higher performing students (generally White and/or more economically advantaged) will have a lower overall AYP target than a school with a smaller number of these students.

persistent factors accounted for about one-half of the variance in mean *gain* scores (using value-added measures between grades 3 and 4) in North Carolina elementary schools of median size (56 students per grade level). Changes in mean *test* scores (e.g., from 4th grade in one year to 4th grade in the next) were measured even more unreliably. Kane and Staiger cite the rarity of repeat winners of rewards for large gain scores as one indicator of the volatility of test scores.

Both studies also showed that the reliability of measurement is related to the size of the group. That is, test scores are less reliable for small schools and/or for subgroups with a small number of students, increasing the chance of misclassification in these cases. Quoting Hill (2000) again:

> A school with some relatively small subgroups might make a real gain with each of its subgroups, but have one or more of the subgroups show inadequate gain because of random error. As groups get smaller, the probability of correct placement goes down, and as the number of such subgroups increases, the probabilities can drop dramatically (p. 18).

The AYP formula contained in S1 is less sensitive to the potential misclassification of subgroups, because it combines the performance of all groups in a school. The Senate bill is still subject, however, to the volatility of school-level year-to-year score changes, particularly in small schools. This volatility is particularly problematic as both bills contain sanctions for schools if they do not improve in one year's time. Schools that fail to make AYP in *any* year are identified for school improvement. If schools fail to improve in the second year, they are subject to corrective action and students have public school choice. A third year of failure can trigger restructuring or reconstitution. These timelines are shorter than those embodied in current law.

One could argue (and I do, later in this paper) that concerns about misclassifying a school (e.g., identifying a school as having not met its AYP target when it fact it has) will vary with the consequences attached to the classification. For example, educators and policymakers may have a greater tolerance for misclassification if AYP data are used only for public reporting. Their tolerance may decrease when failure to meet annual AYP targets triggers labeling (e.g., a school in need of improvement), technical assistance or stronger sanctions, such as school choice or restructuring.

Political feasibility

The combination of short timelines and the proposed AYP definitions may create politically unacceptable situations in many states. First, both bills mandate timelines (maximum of 12 years in HR1; 10 years in S1) for bringing all students to a state's proficient level of academic performance. As of summer 2000, only a handful of states (such as Vermont and Oregon) had established a comparable timeline. Other states either had longer timelines (such as 20 years in Kentucky and California), or had set lower goals, such as bringing only 70% of their students to proficiency (e.g., Maryland or Virginia). The proposed federal timeline may be particularly problematic for those states that have established rigorous standards and implemented challenging assessments. Only 22% of Virginia's schools met their 70% threshold in 2000 (Hess, 2001); a smaller percentage of schools met this mark in Maryland. Only one percent of Massachusetts schools had 80% of more of their students scoring at the proficient or

advanced level on their state assessment. One-quarter of the schools (and 35% of the high schools) had fewer than 20% of their students at this level (MDE, 2001). The percentage of successful schools in all of these states might be lower if subgroup performance is factored in.

Second, large numbers of schools could be identified as needing improvement because they will fail to meet their AYP target in one year, especially if the AYP target includes subgroup performance. The federal AYP requirement could particularly impact higher performing schools in the states. Many states have focused their attention and targeted their resources on persistently low-performing schools, rather than identifying high performing schools that fail to show much growth. Massachusetts, for example, sets, but does not act on, overall improvement targets for schools that have 60% of more of their students at the proficient level. These schools, however, will have to meet AYP targets under either the House or Senate bill. In addition, some of the higher performing schools might not meet subgroup performance targets.

In the past, states that faced politically unacceptable failure rates for either schools or students responded by lowering the stakes, reducing the thresholds to pass, or delaying implementation of the accountability policy (Hess, 2001). In response to public feedback, Virginia policymakers limited the consequences for schools that failed to meet their performance-based accreditation standards to the preparation of school improvement plans (Hess, 2001). Kentucky weakened sanctions for schools that failed to show progress, and reduced the possibility that high performing schools would be identified as needing assistance. When confronted with high failure rates on high school exit examinations, states have lowered passing rates or delayed implementation of the diploma requirement (Fuhrman, Goertz and Duffy, in preparation). States have also been very slow to take corrective actions against schools.

Educational feasibility

States have limited the number of schools they identify for improvement because they lack the fiscal and human resources to provide technical assistance to struggling schools and to oversee corrective actions. Many states rely primarily on federal funds, particularly from Title I, to support program improvement. Thus, one educational issue facing Congress is who will assist the potentially large number of schools identified for school improvement under either bill? Who will fund teacher professional development, new instructional materials and support services for students who are at-risk of not meeting standards?

A second concern is the very short timeline for moving schools from identification to corrective action to restructuring. Research has shown that it takes a long time to turn around failing schools, particularly those with high concentrations of low-performing students (see, for example, Elmore with Burney, 1997; Fullan, 1991). Expecting a school to improve significantly after one year of technical assistance is unrealistic, especially if the improvement target calls for increasing the number of proficient students by seven to ten percent annually. Change is certainly possible. We have many examples of high-poverty, minority schools that educate all of their children to high standards. These schools did not get where they are, however, in two to three years.

Is the Proposed System Flexible?

Both bills prescribe timelines, AYP definitions and types of assistance and corrective action. They give states limited flexibility in the design and implementation of state accountability policy. In some cases, the new ESEA provisions will cause states to take positive steps, such as establishing timelines for getting students to a proficient level of achievement; requiring schools to bring all students to standards, not just 50% or 70% of the student body; and creating incentives for closing the achievement gap. In other cases, federal policy will force states to discard policies that address their structural, demographic and political contexts.

Congress' "one size fits all" approach to education accountability policy will restrict the natural experimentation that has taken place at the state level over the last ten years. The AYP formulas contained in HR1 appear to be based on early state accountability policies, such as those in Texas and Kentucky. The federal provisions do not capture important changes in the Kentucky model (particularly those that account for measurement error), nor do they reflect more complex policies in states like Delaware, Illinois, Massachusetts, Rhode Island and Vermont that consider both performance levels and performance growth, and emphasize assistance over sanctions. The proposed legislation could also invalidate the use of "value-added" accountability systems in states like North Carolina.

In addition, the proposed AYP provisions do not appear to be sensitive to differences in school size or demographics. In particular, the Senate formula will have a different impact on schools that are majority White and economically advantaged than on schools have a heterogeneous population or that are predominately minority and economically disadvantaged.

Is the Proposed System Focused?

The overarching goal of Title I is to level up performance; that is, to bring all or most students to the proficient level on their state standards. The incentive structure should be focused on the continuous growth of all students and subgroups of students. Given limited resources, however, the federal government and the states must give serious consideration to developing differential levels of assistance and sanctions, based on the relative performance and educational capacity of the schools. The most intensive assistance and strongest sanctions should be limited to the lowest-performing schools and students. The House and Senate bills, however, appear to foreclose this kind of differential treatment. Any school that fails to meet its AYP target in a given year is treated the same, regardless of its overall level of performance.

Recommendations

The reauthorization of ESEA presents both a challenge and an opportunity to the President and to the members of Congress. Policymakers need to balance a number of competing and potentially conflicting goals. State accountability systems must set improvement goals that ensure that *all* students meet the state's proficient level of achievement within an appropriate timeframe. That timeframe must be short enough to create and sustain a sense of urgency, yet be realistic. The AYP definitions must ensure continuous progress, yet be flexible. The AYP measures must focus on closing the achievement gap, while providing statistically reliable measures of subgroup performance and growth. Accountability timelines must create incentives to improve practice, yet ensure that schools have sufficient time and resources to

make lasting changes. Schools and school districts must have sufficient resources to help all students reach their state's standards.

In short, Congress must embrace the "flexibility/accountability" trade-off that is at the heart of performance-based accountability systems. The reauthorization of ESEA should include a set of "non-negotiable" criteria for AYP, including the establishment of subgroup performance goals, but not prescriptive formulas. States should have the flexibility to design systems that build on current practice and address their unique structural and demographic characteristics, but must show how these systems will result in the continuous progress of all students and the closing of achievement gaps. Subgroup accountability can be accomplished by extending the proposed timelines; by giving states more flexibility in how and when to identify schools that do not attain AYP and in the kinds and levels of assistance and sanctions they apply to schools; and by using annual public reporting to monitor the progress of all students.

Timelines

1. Congress should set a target date—probably 12 years—by which all students are expected to reach a proficient level of achievement. States, however, should have the opportunity to justify an extension of up to 3 years if they can show that the rigor of their standards is so high and that the baseline level of achievement, including subgroup performance, is so low that the 12-year timeline is not feasible. Mandating a shorter and inflexible timeline may encourage states to set lower proficiency levels (which are not subject to review by the Secretary).

2. Extend the timelines for identifying schools for improvement and for applying sanctions. The timelines in the IASA are appropriate—two years of failing to make AYP to be identified for improvement, and up to three years in improvement status prior to taking corrective action. These extended timelines will address the instability of school-level AYP measures and allow sufficient time to build capacity. Some states, like Kentucky, Massachusetts and Vermont, calculate school ratings every two years rather than annually.

3. Because of test score volatility, require schools to show at least two years of continuous improvement before they come off improvement and/or corrective action status.

4. Do not wait until 2003 to identify, assist and sanction low-performing schools. States are in the process of identifying and assisting schools in need of improvement under IASA. Congress should retain the provision in HR1 that allows states and districts to identify schools that were in improvement status under IASA, and gives them the authority to take corrective action if schools were in school improvement status for at least 2 years prior to the enactment of the new law.

AYP measures

5. Retain the provisions in S1 that require AYP to be defined in a manner "that is statistically valid and reliable" (Sec. 1111(b)(2)(B)(ii)). The Secretary should seek the advise of measurement experts (through NCES or a technical advisory committee) about what

constitutes reasonable but technically sound ways of measuring AYP for schools and for major subgroups within schools.

6. Retain the provision in S1 that allow states to average data over two to three years when calculating a school's AYP (Sec. 1111(b)(2)(D)(ii)).

7. Allow states to incorporate measurement error in their accountability systems. Kentucky, for example, has defined AYP as the distance a school must progress each year in order to reach an index of 100 (or proficient) by 2014. The linear progression that each school must follow is called a "goal line." The state then draws a line that is one standard error of measurement below the goal line (based on schools of comparable size). A school's accountability index must fall below this second line to be identified as needing assistance. A school that falls in between the "goal" and "assistance" line is considered "progressing."

8. Congress should also require states to report the margin of error for any school result and the probability that a school that is identified as not meeting AYP has been misclassified. (See Linn, 2001, and his related publications.)

Subgroup performance

9. Congress should require states to establish separate AYP goals for each major subgroup, and, as required in HR1, *report* progress toward meeting these goals at the state, district and school levels.

10. Schools that meet school-wide AYP but do not meet subgroup AYP in any two-year period, should be issued a *warning* and provided information on programs designed to improve the performance of the particular subgroups. If a school fails to meet its subgroup AYP for three continuous years—and the measures are statistically reliable—then the school should be identified as in need of improvement, regardless of its overall AYP status. This extended timeline should address the greater volatility of subgroup scores, and give higher performing schools time to focus their attention and resources on their achievement gaps.

11. If subgroups are too small to yield statistically reliable AYP measures in particular schools, states should apply AYP formulas that focus on raising the performance of the lowest-performing students. A few states, such as Michigan and Colorado, used "closing the achievement gap" formulas to measure AYP in Title I schools under the IASA. Congress should not substitute "closing the achievement gap" measures for subgroup AYP in the ESEA reauthorization, however. Formulas that target the lowest-performing students, without regard to subgroup membership, could have the same impact as the S 1 definition of AYP. That is, they could enable schools with diverse student bodies to ignore the progress of minority or poor students in the short-term. If states are allowed the option of using a "closing the achievement gap" AYP measure, they should justify this approach to the Secretary, show how it benefits all subgroups, and monitor the progress of subgroups through public reporting.

Differential treatment

Congress should give states the flexibility to develop additional kinds and levels of identification and to tailor and target both technical assistance and corrective action to those schools that need it the most. The goal is to enable states and localities to construct incentives and supports that are appropriate for each school's level of performance and capacity.

12. States should have the authority to issue an "early warning" to those schools that (a) fail to meet their AYP goal, but are performing at or above the state average; or (b) schools that have met their overall AYP goal, but not their subgroup objectives. To use wording from the Massachusetts accountability law, this is a warning that the school "must strengthen its improvement efforts" either generally or for certain student populations. Schools that have been issued a warning might have to develop a plan that shows how they will address their performance shortcomings and will be provided with lists of appropriate resources. The early warning status is a precursor to being designated a school as in need of improvement. Schools that fail to make adequate improvement during this warning period can then be identified "for improvement."

    This "early warning" label will provide a wake-up call to middle performing, complacent schools or to schools that are not adequately serving their minority, disadvantaged or special needs students. It will not, however, carry the same stigma, nor the technical assistance cost, of identifying these schools for improvement.

13. States should have the flexibility to target different kinds and levels of assistance to schools that are identified for school improvement. The most intensive technical assistance should be targeted to low-performing schools that are making the least progress.

14. Corrective action should be focused on the lowest-performing and persistently failing schools. (States and districts currently use this strategy to allocate scarce school improvement resources). Both the House and Senate bills place too many schools into corrective action too quickly, and probably needlessly. Many of these schools need continued technical assistance, but not restructuring. In addition, how many schools can a district or state reconstitute or turn into charter schools in any given year?

District accountability

15. Congress should ensure that districts, as well as schools, are held accountable for student performance and progress. States have been slow to develop district-level systems of accountability. In 1999-2000, only 17 states held districts accountable for student performance or for the performance of their schools. And only five of these states applied the same performance criteria to schools and districts. S1 requires states to apply the same AYP measures at the school, district and state level. HR1 is less clear on this point. Both bills, however, establish procedures for identifying districts in need of improvement and for imposing sanctions.

16. Regardless of the AYP measures a state uses, districts should be held accountable for closing the achievement gap within subgroups. Aggregating subgroup performance to the district

level may provide greater stability to annual measures of subgroup student growth—at least in larger districts.

Adequate resources

17. Congress must fully fund Title I so that states and local school districts have sufficient resources to help poorly performing schools and students. Demanding greater accountability without increasing teacher and school capacity and supporting at-risk students is not only unfair, it won't work.

Table 1. AYP for School 1

Assume that School 1 has 1000 students. Half of the students are White, 25% are African American and 25% students are Hispanic. The school has 10 years to get 100% of the students to the proficient level.

	Enrollment	**# students proficient in Year 1**	**% students proficient in Year 1[5]**	**Annual rate of improvement***
White	500	280	56%	4.4%
African American	250	68	27%	7.3%
Hispanic	250	93	37%	6.3%
Total	1000	441	44%	5.6%

* AYP is: $\frac{100\% - \%\text{students at proficiency in Year 1}}{10}$

[5] These proficiency rates are based on 5th grade reading performance in Maryland in 1999-2000 as reported on the Maryland State Department of Education web site.

Assume that the school has 12 years to get 100% of the students to the proficient level.

	Enrollment	**# students proficient in Year 1**	**% students proficient in Year 1**	**Annual rate of improvement***
White	500	280	56%	3.7%
African American	250	68	27%	6.1%
Hispanic	250	93	37%	5.3%
Total	1000	441	44%	4.7%

* AYP is: $\frac{100\% - \%\text{students at proficiency in Year 1}}{12}$

References

Education Trust (2001, May 8). The White House's proposed school accountability system: Can it work? What would be its likely effects for poor children?

Elmore, R. F., & Rothman, R. (Eds.) (1999). *Testing, teaching and learning: A guide for states and school districts*. A report of the National Research Council's Committee on Title I Testing and Assessment. Washington, DC: National Academy Press.

Elmore, R. F., with the assistance of D. Burney (1997). *Investing in teacher learning: Staff development and instructional improvement in Community School District #2, New York City*. New York and Philadelphia: National Commission on Teaching & America's Future and Consortium for Policy Research in Education.

Fuhrman, S. H., Goertz, M. E., & Duffy, M. C. (in preparation). Slow down, you move too fast: The politics of high stakes accountability.

Fullan, M. G. (1991). *The new meaning of educational change*. Second edition. New York: Teachers College Press.

Goertz, M. E., & Duffy, M. C. (2001). *Assessment and accountability systems in the 50 states: 1999-2000*. RR-046. Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education.

Hess, F. M. (2001). Reform, resistance,…retreat? The predictable politics of accountability in Virginia. Paper presented at the Brookings Conference on Accountability and Its Consequences for Students, Washington, DC. May 15-16, 2001.

Hill, R. (2001). *The reliability of California's API*. Portsmouth, NH: The National Center for the Improvement of Educational Assessment. http://www.nciea.org.

Kane, T. J., and Staiger, D. O. (2001). Volatility in school test scores: Implications for test-based accountability systems. Paper presented at the Brookings Conference on Accountability and Its Consequences for Students, Washington, DC. May 15-16, 2001.

Linn, R. L. (2001). Reporting school quality in standards-based reform. *CRESST Policy Brief 3*. Los Angeles: UCLA. http://www.cse.ucla.edu

Massachusetts Department of Education (MDE) (2001, January 9). 2000 school performance ratings issues for Massachusetts public schools. http://www.doe.mass.edu/ata/ratings00/

National Assessment of Title I, Independent Review Panel (NATI/IRP)(2001). *Improving the odds*. Washington, DC: Council for Basic Education.

U. S. Department of Education, Planning and Evaluation Service (2001). *High standards for all students: A report from the National Assessment of Title I on progress and challenges since the 1994 reauthorization.* Washington, DC: Author.